I. Introduction.

Cancer is a common, complex, and frequently fatal disorder, in which the mechanisms that normally govern the growth and functions of our cells go awry. How can this disease be mastered when our knowledge of the behavior of normal cells remains rudimentary? We contend that the best hope lies in a newly flourishing enterprise, one significantly nourished by our own work, that applies the techniques of molecular biology to a relatively small set of genes recently implicated in the development of cancer ("oncogenes").

The conviction that cancer is a disease involving structural alterations (mutations) of genes has its roots in a variety of clinical and experimental observations. Carcinogens are frequently mutagens; certain neoplastic diseases display patterns of inheritance resembling genetic disorders; cancer cells often manifest gross distortions of their chromosomes; and the phenotype of a cancer cell has a stability evocative of genetic change. However, to the many experimentalists entering cancer research as we did 10 to 20 years ago, the most striking fact was that the permanent addition of one or a few viral genes to a normal cell could convert it to a cancer cell. Compared with the daunting complexity of the vertebrate genome, the simplicity of certain tumor viruses (the polyomaviruses and retroviruses) seemed refreshingly approachable. Although the study of such viruses and affected cells from chickens and rodents seemed at first to represent an expedient compromise with Nature, there have been rich and unexpected rewards, particularly for those of us working with retroviruses. The many different oncogenes found to be carried by retroviruses (v-onc's) have proven to be derived from normal cellular genes (known as cellular oncogenes [c-onc's] or proto-oncogenes); moreover, the cellular oncogenes show many signs of being the targets for various mutational events that lead to cancer in animals and man.

Our laboratory first unveiled cellular oncogenes by finding that the DNA of birds and mammals contains homologues of v-src, the oncogene of Rous sarcoma virus. This and subsequent findings prompted us to propose that v-src arose by the capture of a cellular gene (c-src) by a pre-existing retrovirus without an oncogene; that c-src is a gene highly conserved in Nature and vital to normal cells; and that c-src (or genes like it) might figure in the genesis of many cancers, regardless of precipitating cause.

In the decade that followed, several developments enlarged the repertoire of cellular oncogenes and strengthened the argument that they are involved in many forms of cancer. (i) Definition of the composition and origin of retroviral oncogenes other than v-src, here and elsewhere, uncovered many new cellular oncogenes, nearly twenty at last count; as the number grew, so did the variety of biochemical mechanisms for inciting neoplastic growth and the prospects for perceiving mechanisms relevant to human cancer. (ii) The study of retroviruses lacking their own oncogenes provided the first direct evidence that known cellular oncogenes could participate in carcinogenesis, as the targets for activation by viral insertional mutations. This mechanism also affords a novel means to search for new oncogenes. (iii) DNA-mediated gene transfer into rodent cells has uncovered active oncogenes in human and other tumors. Again cellular oncogenes previously identified by their homology with retroviral oncogenes have frequently been implicated, and an explicit definition of carcinogenic change at the nucleotide level has been possible for the first time. (iv) Cellular oncogenes or their close relatives have been encountered amidst chromosomal

anomalies with surprising frequency, the genes sometimes amplified in numbers and extent of expression and sometimes translocated from one chromosome to another.

Such findings not only validate the study of retroviral oncogenes as models for the biochemical basis of human cancer, they also invite a direct assault on an apparently limited set of cellular genes important in the creation of a cancer cell, irrespective of external cause. The questions we address in our work and this proposal spring directly from this perspective. What is the full roster of cellular genes instrumental in carcinogenesis? What are the normal functions of these genes? How do their oncogenic homologues---either mutant alleles in tumors or transduced derivatives in viral genomes---differ structurally and functionally from their normal progenitors? What are the biochemical consequences of oncogene activation and how do those consequences lead to the loss of growth control and the other changes that typify a cancer cell? Might we be able to use the emerging knowledge of oncogenes to begin the rational design of strategies for the control of cancer?

II. Expanding the roster of oncogenes.

The size of the complete repertoire of cellular oncogenes is unknown. Retroviruses have brought to light at least twenty such genes, and more appear to be in the offing. For example, two leukemia viruses (E26 and MH-2) under study in our laboratory harbor as yet uncharacterized genetic loci that may be oncogenes. Several oncogenes have also been added to the list by the study of insertion mutations, oncogenic DNA, and chromosomal rearrangements.

We are placing special emphasis upon the use of insertion mutations to identify new oncogenes. The value of this approach has been substantiated---and given rise to yet more experimental opportunities described below---in studies of avian B cell lymphoma and mouse mammary carcinoma. In the lymphomas, a cellular oncogene (c-myc), previously discovered here by its homology with the oncogene of MC-29 virus, was found to be activated in virtually all tumors by adjacent insertions of avian leukosis virus DNA. This finding presaged other kinds of evidence for the involvement of c-myc in human and murine tumors: amplifications of the c-myc gene and chromosomal translocations that join c-myc to immunoglobulin loci. In mouse mammary tumor virus (MMTV)-induced carcinomas, we have traced proviral DNA to the site of insertion mutations and thereby identified a cellular gene (called int-1) that is activated by the insertions. This gene, like other putative oncogenes, has been highly conserved during evolution, but expression of it has been observed to date only in mouse mammary tumors bearing nearby proviral DNA. We now seek to know the function of this gene (see below), and whether it figures in non-viral carcinogenesis in human beings and other animals. We are also attempting to discover new oncogenes in two other contexts in which insertion mutations may be operative: in nephroblastomas induced in chickens by myeloblastosis-associated virus, and in primary hepatic carcinomas (PHCs) associated with infection by hepatitis B virus (HBV). The study of nephroblastomas may have special rewards: no gene has yet been implicated in human renal cancer, though several chromosomal abnormalities have been described; the precedent of c-myc sugggests that targets for insertion mutations may also be involved in chromosomal rearrangements. PHC has a significance that is self-evident: this disease is among the most common fatal cancers of man worldwide. In both contexts we seek chromosomal domains that are physically and

transcriptionally altered by viral DNA. Previously known cellular oncogenes seem thus far not to serve as targets for insertion mutation in these tumors, but in one hepatoma we have an interesting lead: an integrated unit of HBV DNA and flanking cellular DNA is many-fold amplified.

Similar considerations influence the search for oncogenes in tumors without apparent viral cause. For example, the thought that gene amplification might bring cellular oncogenes into the tumorigenic scheme led to our recent discovery of a gene we call N-myc - a distant kin to c-myc that is amplified in human neuroblastomas. Our study of neuroblastoma was motivated by karyological evidence that the tumor frequently, if not inevitably, harbors a domain of amplified DNA. Searching within this domain for representatives of known oncogenes, we encountered the previously unrecognized N-myc. We now know that N-myc is present throughout the vertebrate phyla (as expected for a candidate cellular oncogene); that the haploid genome of human cells contains a single copy of N-myc, situated on the short arm of chromosome 2; and that amplification of N-myc may be a molecular marker for neuroblastoma, perhaps even an etiological factor in the genesis of the tumor. The rationale that engendered the discovery of N-myc seems worthy of pursuit in the numerous other examples of human tumors presenting with karyological evidence of gene amplification prior to therapy.

We wish to note here an aspect of our work that may appear not to be directly related to oncogenes and their functions: the study of mechanisms by which the retroviruses and hepatitis B viruses replicate. In reality, this work has provided an important theoretical base for many of our excursions among oncogenes; in particular, it inspired the search for retroviral insertion mutations that activate cellular oncogenes. We have a continuing commitment to study retroviral and host factors important for integrative recombination; the transposition of genetic elements structurally related to retroviral proviruses in Drosophila; and the genetic program and replicative strategy of hepatitis B viruses in lower mammals.

III. The functions of retroviral oncogenes.

The oncogenes carried by highly tumorigenic retroviruses are among the most potent and experimentally malleable carcinogenic reagents known. Within hours after infection---or within minutes after temperature shift of cells infected with thermosensitive mutants---normal cells can be converted to tumorigenic cells through the action of modest amounts of a single protein. Furthermore, work here and elsewhere has sketched some preliminary versions of how such proteins act; for example, several of the v-onc proteins exhibit protein kinase activity specific for tyrosine and reside in the plasma membrane. Future challenges include the identification of (i) the functionally-significant cellular targets for tyrosine kinases; (ii) the structural features of the transforming proteins important to their localization and activity; and (iii) the biochemical properties of those oncogenic proteins that are not tyrosine kinases. The ultimate goal is to understand in detail the apparently varied oncogenic mechanisms used by viral genes whose relevance to human cancer can now hardly be questioned.

We continue to place special emphasis upon the src gene of Rous sarcoma virus and its product, a 60,000 dalton phosphoprotein ($pp60^{v\text{-}src}$) that phosphorylates tyrosine residues in many proteins. The power of a joint genetic and

biochemical approach is manifest here: we are isolating and engineering new src mutants to probe (i) the basis and significance of the tyrosine kinase activity, (ii) the interaction of pp60$^{v\text{-}src}$ with various cellular proteins, (iii) the importance of serine and tyrosine phosphorylations within pp60$^{v\text{-}src}$, (iv) the tertiary structure and membrane attachment of the protein, and (v) our recent finding that src may encode a second protein in an alternative reading frame. In addition, we are attempting to isolate host mutants that fail to respond to v-src in hopes of identifying: host factors required to mobilize pp60$^{v\text{-}src}$, immediate targets for phosphorylation, and other components in the transformation pathway. We have mutants in hand that display only a portion of the transformed phenotype and some that are competent to affect some host cells but not others; these too will be useful in a dissection of transformation by src. Recently we have placed the v-src gene under hormonal regulation (by linking it to the glucocorticoid-sensitive promoter from the mouse mammary tumor virus); this maneuver has produced another system ripe for biochemical exploitation, since we can now reproducibly modulate the phenotype by making relatively small adjustments in the dose of pp60$^{v\text{-}src}$ and its kinase activity.

If we are to realize the full benefits of genetic analysis, we must learn much more of the biochemistry of pp60$^{v\text{-}src}$. Important objectives include: (i) more effective purification of the protein in bulk, perhaps by exploiting production in bacterial hosts; (ii) detailed characterization of the enzymatic reaction catalyzed by pp60$^{v\text{-}src}$ and comparisons with better known protein kinases; (iii) identification of the amino acid sequences that define preferred sites for phosphorylation by pp60$^{v\text{-}src}$; (iv) the design and synthesis of peptides that represent model substrates and potential inhibitors of the kinase activity; (v) the use of inhibitors to obtain more decisive evidence that tyrosine phosphorylation is responsible for the tumorigenic capacity of pp60$^{v\text{-}src}$; and (vi) the development of more effective strategies for the biochemical recognition of cellular proteins phosphorylated by the enzyme (approaches could include affinity chromatography on purified pp60$^{v\text{-}src}$ and antisera to both phosphotyrosine and peptide substrates). As the mechanistic analyses of src advance, principles by which the effects of the gene on cells might be reversed should emerge, presenting us with new opportunities to interdict malignant growth.

There is little in the design of our studies on v-src that cannot also be applied to the other viral oncogenes that have caught our fancy (v-fps, v-myc, v-erb-A, v-erb-B, and v-myb). Although only one of these (v-fps) is a protein kinase, genetic analysis of both viral gene and cellular response, purification of the protein products of the genes, and biochemical quests for the cellular macromolecules with which the transforming proteins may interact will nevertheless be central to our explorations of how these varied genes transform cells to neoplastic growth. We believe that it is essential to study a multiplicity of oncogenes in order to achieve a fair sampling of the biochemical mechanisms that can underlie tumorigenesis, and to learn how these mechanisms affect cells of different embryological lineages.

IV. The normal and carcinogenic functions of cellular oncogenes.

The findings that implicate specific cellular genes in tumorigenesis---whether the genes were encountered first in retroviral genomes (e.g., src), at sites of insertion mutations (e.g., int-1), or in amplified DNA (e.g., N-myc) ---raise two obvious but unanswered questions: what are the normal functions of

those genes? and how do alterations of them contribute to disease? The clues at hand are inspirational but not decisive: many of these genes are strikingly conserved throughout evolution, arguing that they are required for some fundamental life process; variations in expression of these genes among tissues or during embryogenesis and the affiliation of certain oncogenes with certain types of tumors suggest roles in development or differentiation; and comparisons of viral oncogenes (or mutant cellular oncogenes) with their progenitors reveal differences in both the structure and abundance of products, implying that both qualitative and quantitative factors may operate.

The specific objects of our attention are the cellular genes homologous to the viral oncogenes discussed in the preceding section (src, myb, myc, fps, erb-A, erb-B); genes uncovered as the targets of insertion mutation (int-1, as well as myc and erb-B, and perhaps others yet to come); and genes involved in chromosomal rearrangements (N-myc and c-myc, principally). The level at which we can approach the overriding questions must vary in each case. For genes of which we are largely ignorant (e.g., int-1 and N-myc), the primary issues are a more complete structural definition of the genes and the identification of the protein products. In cases for which proteins have been identified (e.g., c-src and c-myc), the immediate issues are more sophisticated: discovering the relevant biochemical properties of the gene products and defining the genetic and biochemical distinctions between normal and altered proteins.

We and others have wagered that cellular oncogenes may play important roles in the growth and development of normal organisms. This is an exciting prospect, because biological scientists have previously had little genetic purchase on cell division and differentiation in higher eukaryotes. But how are the roles of cellular oncogenes in normal cells to be sought? In the belief that genetic strategies are likely to be most telling, we have turned to the two experimental systems that offer the most facile access to the genetics of eukaryotes: yeast (Saccharomyces cerevesiae) and the fruit fly (Drosophila melanogaster).

Genes related to many (perhaps all) of the cellular oncogenes of birds and mammals can be found in Drosophila. The great store of information accumulated from classical genetic analysis, as well as recent remarkable progress with techniques to modify the genome of Drosophila, should make it possible to ask directly how cellular oncogenes might contribute to the growth and development of the fly. The prospects for this work can be demonstrated by our experience to date with Drosophila src: the identity of the gene has been documented by nucleotide sequencing; the gene has been mapped to position 64B on chromosome 3; a tyrosine-specific protein kinase apparently encoded by the src locus has been identified in tissues of Drosophila; and expression of the locus has been shown to fluctuate dramatically during the course of Drosophila embryogenesis. These findings prepare the way for a mutational analysis of Drosophila src and have emboldened us to proceed with isolates of other cellular oncogenes, including fps, ras, myc, N-myc and int-1.

Although it is conceivable that cellular oncogenes would serve different physiological purposes in yeast and in metazoan organisms, the ease and precision with which the genome of yeast can be manipulated call for exploitation whenever possible. Cautious excitement has therefore greeted the discovery here and elsewhere that yeast may harbor recognizable members of the cellular oncogene family. Candidates to date include src, fps, myc, mos and ras, and the

search is on for others. Once the identity of any of these genes has been validated, targeted mutagenesis and other genetic manipulations should quickly reveal whether the gene is essential to yeast and what its physiological purpose might be.

Genes can now be introduced into the genome of mice by microinjection of single-cell embryos. We hope eventually to use this remarkable technology in our search for the normal and pathogenic functions of cellular oncogenes, and we have begun to develop the necessary reagents. But there are presently serious constraints on the utility of the procedure, since it has not yet proven possible to dictate whether, when or where the implanted gene will be expressed. We will wait for relief of these limitations before pursuing the creation of transgenote mice for our purposes.

Views of how the growth of normal and neoplastic cells is governed have merged dramatically in the study of polypeptide growth factors. At least one of these, platelet-derived growth factor (PDGF), is a close genetic relative of the protein encoded by the viral oncogene v-sis; several (PDGF, epidermial growth factor [EGF], and perhaps insulin) elicit phosphorylation of tyrosine in cellular proteins, possibly by activating tyrosine-specific kinase activity residing on the receptors to which the growth factors bind; and illicit production of growth factors may be a central anomaly in many forms of cancer cells. Anticipating (and provoked by) these associations, we have initiated two lines of attack designed to explicate the cellular response to growth factors. First, we have sought and found modifications in the structure and function of the protein encoded by the cellular src gene during the early cellular response to PDGF. The nature of the modifications and their role in the response to PDGF are under study. Second, we have begun the molecular cloning of genes that encode the receptors for EGF and PDGF. The cloned genes will reveal valuable details of receptor structure. More importantly, the clones can be manipulated in vitro and then transplanted back into living cells in order to dissect the role of the receptors in the response to their ligands.

How large is the role of cellular oncogenes in carcinogenesis? What changes must be wrought in these genes to make them pathogenic? How might these changes affect the functions of the genes? We seek the answers to these questions in a multifaceted approach that lies at the core of our research program, and that employs all of the oncogenes now under our scrutiny.

(i) The protein products of these genes must be identified, produced in quantity, characterised in detail and compared to their pathogenic kin. In these efforts we are assisted by the remarkable progress that has occurred in the cloning and expression of eukaryotic genes in bacteria and in the production of poly- and monoclonal antisera with antigens synthesized chemically or in bacteria.

(ii) The structural details of cellular and viral oncogenes must be elaborated and compared in a search for features that can modify function.

(iii) The carcinogenic capacity of cellular oncogenes needs to be tested persuasively in cell culture and in animals, using viral vectors and other means of gene-transfer. Here the dividends of pursuing retroviral replication are again apparent: various versions of cellular oncogenes - including those such as int-1 and N-myc, not previously found in retroviral genomes - can be harnessed

to retroviral elements and delivered throughout a host animal or to all cells in a culture.

(iv) The genes can be manipulated in vitro to describe the variety of structural alterations that can produce or enhance pathogenicity. The starting point in such efforts is the genesis of hybrid genes that combine domains from viral oncogenes and their cellular progenitors, but manipulation of the cellular genes alone should also be revealing, especially when retrieved from normal and cancerous cells. The functional consequences of the manipulations can be tested by gene-transfer in culture and in animals.

(v) Further evidence to implicate cellular oncogenes in the genesis of human tumors must be sought. All of the strategies we have learned before, and others yet to come, can be employed in this search: pursue new oncogenes identified by insertional mutagenesis in the appropriate human tumors (for example, int-1 in carcinoma of the breast, and the presently unknown genes we hope to find in hepatic and renal carcinomas); exploit the amplification of DNA in human tumors to implicate known cellular oncogenes in tumorigenesis and to find potentially novel oncogenes; continue our search for inordinate expression of known cellular oncogenes in diverse samplings of human tumors; use established procedures for gene-transfer to seek active oncogenes in the tumors we study; use viral vectors to develop more efficient and more comprehensive strategies for the rescue of tumorigenic genes from human neoplasia; and explore the use of gene-transfer and somatic cell genetics for the detection of faulty regulatory devices that might contribute to neoplastic growth but that would not be counted as members of the oncogenic family.

V. Epilogue.

We make this proposal with two firm convictions: that cancer is at its heart a genetic malady, and that retroviruses have provided us with our best grip on the malady. If there is a common genetic substrate on which various carcinogens act, cellular oncogenes are likely to be part of that substrate. If we can unravel the means by which these genes act, we should be able to perceive at least the outlines of a rational design for the prevention and cure of cancer.